

Communication in Private-Information Models: Theory, Computation, and an Application

Edward S. Prescott¹
Federal Reserve Bank of Richmond

February 23, 2001

¹This paper is based on Chapter 2 of my August 1995, University of Chicago Ph.D. dissertation. I would like to thank my dissertation advisor, Rob Townsend, for his advice, support, and helpful conversations. I would also like to thank George Hall, Charlie Kahn, Mark Penno, Benjamin Sacks, and José Scheinkman, for their helpful comments. Versions of this paper were presented at the Federal Reserve Bank of Richmond, Queen's University, and the Universities of Chicago, Iowa, Rochester, and Western Ontario. My gratitude is extended to the Jacob K. Javits Fellowship Program of the U.S. Department of Education and the University of Chicago for financial support. Finally, I would like to thank the Mellon Foundation for supporting my field research in Aurepalle, India and ICRISAT for their assistance with it. All errors in this paper are mine. The views expressed in this paper are not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

Abstract

Communication and no-communication versions of a two-stage principal-agent model are compared. The models contain a risk-averse agent and two sources of private information, a shock to preferences followed by a productive action. Both models are formulated as linear programs, which are then used to compute solutions to examples. In simple cases, a sufficient condition for communication to be valuable is derived. In these cases, communication improves risk-sharing in bad states of the world. In more complicated cases, computed examples demonstrate how communication may also alter labor supply. Further examples demonstrate how action and consumption lotteries may separate agents by their shock. Survey evidence and field work on contractual arrangements between farmers and landowners in an Indian village is used to motivate the class of models. Suggestive evidence on communication from the same sources is also described.

Keywords: Private information, communication

1 Introduction

This paper investigates the effects of communication in a class of multi-stage principal-agent models. It provides analytical results for simple cases, develops numerical methods to investigate more complicated cases, and reports on field research that attempted to measure interim shocks and communication in tenant-landowner agricultural contracts.

The environment consists of a risk-averse agent who receives a shock to his marginal disutility of labor. After receiving the preference shock, the agent takes a hidden action which affects the probability distribution of a publicly-observed output. Three models are considered: one where the preference shock is public information, one where it is hidden information but the agent may send unverifiable reports to the principal, and one where the shock is also hidden information but the agent may not communicate to the principal at all. Sending a report is equivalent to selecting from a menu of contracts and is potentially beneficial because compensation schedules can be tailored to the shock the agent receives. Comparison of the latter two models is the main focus of this paper.

The main analytical result demonstrates how communication alters compensation schedules in order to improve consumption risk sharing. In the finite-action case, a sufficient condition for communication to be valuable is that the optimal solution to the no-communication model be characterized by an incentive constraint that does not bind on the type with the high marginal disutility of effort. For reasons similar to that of Penno (1984), the communication contract improves welfare because it can include more insurance for this type without negatively affecting incentives. In contrast, if this incentive constraint binds then communication may or may not be valuable.

These results contrast with those of Melumad and Reichelstein (1989) and Demougin (1989). In an environment with a risk-neutral agent, they find that an invertible technology is a sufficient condition for communication to be valueless. With an invertible technology,

no-communication contracts span the space of expected consumptions and thus can duplicate the incentives faced by a risk-neutral agent in any communication contract.

Other analytical results in the literature contain results for the case where the hidden information is fully revealed by specific outputs. Christensen (1981) describes consumption allocations as a function of outputs and Lagrangian multipliers. He argues that communication is valuable if there is a positive probability that hidden information is revealed ex post to the principal. Christensen's claim is proven in Baiman and Evans (1983). Suh (1990) proves that communication is valuable and characterizes allocations for a simple environment where, again, the hidden information is fully revealed by specific outputs.

The results in the literature depend on assumptions like risk neutrality or full revelation of the hidden information. To obtain further results, this paper develops numerical methods. In particular, it provides linear programs which can be solved to find optimal allocations for arbitrary specifications of preferences and technology. Solutions to examples are computed and then used for inter-model comparisons. To accomplish this, sets of preference shocks, actions, outputs, and consumptions are gridded into finite numbers of elements. Randomization is allowed in contractual terms.

In developing these numerical methods, the following results are provided. First, the communication model is formulated in terms of interim utilities. These interim utilities are a more efficient way to account for off-equilibrium strategies than the standard formulation provided in Myerson (1982). The result is a linear program with a much smaller number of constraints than in the standard formulation. Second, a revelation-principle like argument is provided to formulate the no-communication model as a linear program.

Next, the linear programs for all three models are used to solve numerical examples. The examples are used to address various features of the models. An example with changing labor supply is solved. Another example, where action lotteries are used to help separate types is provided. The final example studies a case where the shock is to the agent's risk

aversion. In this example, consumption lotteries are used to distinguish agents by their shock.

While this paper is primarily a theoretical and computational exercise, the choice of models is motivated by data collected from surveys of tenant-landowner contracts from an Indian village. The data is not used to test the models in a statistical sense. The data is closer to field research as often used by anthropologists and sociologists. It is used to argue that multi-stage principal-agent models capture important features of economic environments and that communication occurs, when feasible, between the principal and the agent.

Finally, it should also be noted that these class of models are applicable to a variety of other environments. They can be used to study managerial accounting or reporting from subdivisions to distant corporate headquarters. They are also applicable to bank regulation. In an attempt to more finely identify the riskiness of banks' assets, bank capital regulation has moved towards requiring banks to report to regulators the risks they face and basing capital requirements on these reports. Since it is difficult for regulators to verify these reports, bank regulators face similar problems to those discussed in this paper. See, for example, Rochet (1999).

2 Data

This section uses data on tenant-landowner relationships to argue that multi-stage models capture important features of agricultural production. The data comes from the results of a set of questionnaires described in Townsend and Mueller (1998).¹ They regularly surveyed eight of thirteen identified cropping groups in the Indian ICRISAT village of Aurepalle for the duration of the 1990 farming season.² These cropping groups are partnerships

¹Additional information on the cropping groups was gathered from interviews with villagers conducted by the author during August 1994.

²For more information on Aurepalle and the other ICRISAT villages see Walker and Ryan (1990).

of 2-5 members who jointly farm plots of land. The surveys collected information on a variety of topics. For the purposes of this paper, only the evidence on interim shocks and communication is used and the cropping group is treated as a single individual. This latter assumption is driven by the nature of the data, though there are conditions, described in Prescott and Townsend (1999), under which the group member's preferences can be aggregated into that of a single individual.

2.1 Interim shocks

Agricultural production requires several sequential operations. Land must be plowed and sowed. Fertilizer and pesticides may be applied. At different times during the season weeding may be required. Finally, the output must be harvested at the end of the season. The timing of these operations is dependent on weather and soil conditions. For example, according to respondents delays in planting seed after the rains can lead to substantial declines in crop yields.

The surveys provide information on the frequency and reasons for delays in performing operations and applying crop inputs. While there are only eight different tenant-landowner relationships in the data set, each group performed numerous operations on their plots. The surveys intensively monitored only a subset of each group's plots but for these monitored plots the eight groups performed 127 operations and applied 36 inputs, numbers which are large enough to provide an idea about the frequency of shocks.

Tenants responses to the operation questionnaires indicate that 28 of the 127 operations performed on the surveyed plots were delayed. Not surprisingly, weather shocks were important. For example, shocks attributable to nature, such as the unavailability of well water, contributed to 19 of the 28 delays. Another important source of shocks, was the amount of time available for group work. For example, the unavailability of bullocks (needed for draft power) contributed to eight of the operation delays. Bullocks could be unavailable

because they were sick or occupied with tenant's own crop work. Also, work on other crops was cited as a reason for six of the delays, though it is not known whether this delay was due to time spent working on other group plots or on own plots.

Application of crop inputs such as seed or fertilizer were also delayed 10 out of 36 times according to tenants. The reasons for these delays are similar to those that caused delays in operations. Shocks that can be attributed to nature, such as rain or soil moisture, contributed to delays in 6 of the 10 inputs. Shocks to tenants, such as bullock unavailability, other crop work, and time constraints, contributed to delays on 4 of the 10 delayed input applications.

Shocks may not only affect the timing of operations and labor but they also may affect the quantity of labor supplied. Some indirect evidence is provided by the surveys. Questions were asked, *conditional* on the group working, whether any members missed work.³ With respect to groups this lowers the amount of labor available on a particular day. In this paper, where there is a single agent, this data should be considered as indicative of some of the shocks which may affect individual's decision to work on the landowner's plot of land.

The data in the first row of Table 1 reports for each group the number of days per group member absent from group operations. Excluded are days missed from irrigation operations.⁴ There is variation across groups. The number of days per group member missed by a group varied from three to sixteen, and the average number of days missed per person among all groups is six. One indication of whether the reported number of absences affects output is to compare absences with respondent's answers at the beginning of the season about the acceptable number of absences. These responses varied from one to seven days. Groups 2, 10, and 13 seem to have a large number of absences.⁵

³Unfortunately, the surveys do not provide evidence on whether shocks influenced the decision of a group to work the group plot or not.

⁴Irrigation absences are left out because this operation is frequently done by a single individual rather than the group.

⁵It should be noted that group 2, which has the second-largest number of absences, expelled a group

Group No.	2	6	7	8	9	10	12	13
Days	13.5	3	3	3.2	5	16.5	4.5	7
Bullock Days	7	2	1.33	1.8	2.66	8	2.5	8

Table 1: Average Number of Days Tenants Are Absent from Group Operations.

Several reasons were cited for the absences. Sickness accounted for one-third, crop work was cited for one-fifth, and non-crop work was cited for approximately one-fourth of the absences. One-fifth of the absences did not fit into any of the listed categories, and hence little is know of their cause.

A further source of information, and a partial check on the absence data, is the average number of days missed from bullock operations by each group. These data are listed in the second row of Table 1 and show that from one to eight days per group member were missed by groups. The average number of days missed among all individuals in the sample is 3.5. The most common reason, cited about one-half of the time, was that the bullock driver was not available. One-fifth of the time, the bullocks were being used for own crop work. Most other responses indicated that bullocks were sick, too weak or that the driver was sick.⁶

The previously described shocks can be classified into two categories. The first are production shocks such as low well water levels or not enough rain. The other class of shocks, and the one which will be studied, are those which affect the value of a farmer's time. First, farmers and their bullocks can get sick and this affects their time endowment or at least their desire to work. Second, if outside opportunities are valuable, such as when work needs to be done on a farmer's own plot, then that farmer has an incentive to work on his own plot rather than the group's plot.

member towards the end of season for missing too much work.

⁶Group 13 indicates more absences from bullock operations than all operations. This could be due to imperfect recall by the respondents.

The surveys also collected some information on communication between tenants and their landowners (as well as their financiers). This information is described and discussed at the end of the paper.

3 Models

The key features of the environment are the multiple stages required for production and the variety of shocks. Because modeling all the stages is not practical, a class of models with two stages is developed. In the first stage there is a shock which affects the agents time endowment or equivalently their marginal disutility of effort. Shocks described earlier such as the health of farmer or his bullocks clearly have this effect. Other shocks, such as those which make working off of the landowner's plot desirable, can also be modeled using these preferences if effort is considered to be only that which is worked on the landowner's plot. The idea is that valuable outside opportunities make it less desirable to work on the landowner's plot. The rest of the shocks described by the data, namely the nature shocks such as rainfall, are best modeled as affecting the production technology rather than preferences. This is an interesting and closely related problem but it is not the focus of the paper.

Three models in this two-stage class are developed. For all three models preferences, technologies, and timing structures are identical. However, assumptions about information and communication differ in ways that will be described shortly. At time $t = 0$ the principal and the agent agree to a contract. After agreement, there is a shock $\theta \in \Theta$ to the agent's preferences. Upon observing the shock θ the agent sends report $r \in R$ to the principal. The principal replies to the report with message $m \in M$. Next, the agent takes action $a \in A$. Finally, output $q \in Q$ is observed, and the principal gives to the agent consumption $c \in C$, as specified in the contract. Specified in the contract are two rules: a randomizer

which sends messages (possibly randomly) as a function of the report (or shock in the case of the public-shock model), and a compensation schedule which may depend on outputs, reports (or shock in the case of the public-shock model), and messages.

Sets C, Q, A , and Θ contain finite numbers of elements. The size of sets R and M are not fixed but depend on the model. Preferences for the agent are defined over the space of lotteries over $C \times A \times \Theta$, and preferences for the principal are defined over the space of lotteries over $C \times Q$. The agent's and principal's ex ante utilities are $\int_{c,a,\theta} \text{prob}(c, a, \theta) U(c, \theta, a)$, and $\int_{c,q} \text{prob}(c, q) (q - c)$, respectively. The production function, denoted $p(q|a)$, describes the probability of output q given action a is taken. Assume that $p(q|a) > 0, \forall a \in A, q \in Q$. The shock θ is random and drawn from the probability distribution $\text{pr}(\theta)$. Reports r and messages m have no direct effect on preferences or technology.

In all three models the action is hidden, known only by the agent. Model P1, the public-shock model, assumes that the shock θ is public so reports are irrelevant for this model. The other two models assume that the shock is observed only by the agent but they differ in the agent's reporting possibilities. Model P2, the communication model, puts no restrictions on the size of R , though by the revelation principle we will be able to restrict ourselves to a reporting set equal to the set of shocks Θ . Model P3, the no-communication model, forbids communication. Communication is forbidden by restricting the size of R to one element – the act of physically not sending a report. Finally, there is no restriction on the size of the set of messages which the principal may send. Figure 1 summarizes the three models. These assumptions mean that consumption schedules in Models P1 and Model P2 may depend on the shock θ but the consumption schedule in Model P3 may not.

Working in the space of lotteries over a finite grid of variables is not the standard method of analyzing principal-agent models. There are two reasons for using this space. First, randomization can be valuable in mechanism design problems. As will be demonstrated later, randomization may relax incentive constraints as well as convexify constraint sets.

Contract reached	Shock $\theta \in \Theta$	Agent sends report $r \in R$	Principal replies with $m \in M$	Agent takes action $a \in A$	Output $q \in Q$ observed	Agent receives $c \in C$
P1: Public- θ model	Public	Size of R is irrelevant	M is as big as necessary	Private to agent	Public	
P2: Communication model	Private to agent	R is as big as necessary	“	“	“	
P3: No Communication model	Private to agent	R has only one element	“	“	“	

Figure 1: Timeline and assumptions on the three models

Second, randomization combined with the finite grids means the problem becomes a linear program, which can be used to compute solutions for arbitrary preferences and technologies. See Myerson (1982), or E.C. Prescott and Townsend (1984).

The next sections develop each model into a linear program. In the interests of brevity and clarity, the revelation principle is invoked and in the main body of the text only the truth-telling and obedience games and their representation as linear programs are discussed. For Model P2, the communication model, an alternative formulation with interim utilities is also provided. A revelation principle proof of Model P3, which is not a standard model, is contained in the Appendix. The other two models are special cases of Myerson (1982) so the revelation principle proof contained there applies. Discussion of the properties of the models is deferred until the section on results.

3.1 Model P1

This model assumes that shock θ is public information. Some preference shocks, such as severe illness or injury, are publicly observable to other residents in the village. The problem for the principal is to design state-contingent allocations which satisfy the moral hazard problem in that state and yet still insure the agent over the states.

Reports are irrelevant in this model because θ is public; thus, all reference to set R is dropped. By the revelation principle the message space can be restricted to the set A . The principal chooses two rules in the contract: a randomizer which, given θ , recommends an action $a \in A$ with probability $\pi(a|\theta)$, and a consumption schedule which specifies the probability of consumption c , given output q , message a , and shock θ . Denote this rule $\pi(c|q, a, \theta)$ and note that degenerate rules are feasible with this notation.

One type of constraint which must be satisfied by a contract for it to be feasible are the incentive constraints.

Definition 1 *An incentive-compatible allocation for Model P1 is a randomizer $\pi(a|\theta)$, a consumption schedule $\pi(c|q, a, \theta)$, and a strategy $a^*(a)$, such that, given the randomizer, and the consumption schedule, $a = a^*(a)$ satisfies*

$$\forall \theta, \quad \times_{c,q} \pi(c|q, a, \theta)p(q|a)U(c, \theta, a) \geq \times_{c,q} \pi(c|q, a, \theta)p(q|\hat{a})U(c, \theta, \hat{a}) \quad \forall a, \hat{a} \in A. \quad (1)$$

In the principal-agent framework, the principal chooses a feasible allocation which maximizes his objective function and satisfies the agent's participation constraint. The Pareto frontier is obtained by solving the following program for each feasible value of the agent's reservation utility level \bar{U} .

$$\begin{aligned} & \max_{\pi(c|q,a,\theta) \geq 0, \pi(a|\theta) \geq 0} \times_{c,q,a,\theta} \pi(c|q, a, \theta)p(q|a)\pi(a|\theta)pr(\theta)(q - c) \\ \text{s.t.} \quad & \text{P} \quad \times_{c,q,a,\theta} \pi(c|q, a, \theta)p(q|a)\pi(a|\theta)pr(\theta)U(c, \theta, a) \geq \bar{U}, \end{aligned} \quad (2)$$

(1),

$$\forall \theta, \text{P}_a \pi(a|\theta) = 1, \quad \forall q, a, \theta, \text{P}_c \pi(c|q, a, \theta) = 1, \quad (3)$$

where equation (2) is the agent's participation constraint, and equations (3) ensure that the choice variables are probability measures.

Transformation of the above program into a linear program requires several steps. First, substitute $\pi(c, q, a|\theta) = \pi(c|q, a, \theta)p(q|a)\pi(a|\theta)$ into the objective function and the participation constraint. Next, make the incentive constraints, equation (1), linear by using the transformation,

$$\pi(c|q, a, \theta)p(q|\hat{a}) = \pi(c, q|a, \theta) \frac{p(q|\hat{a})}{p(q|a)},$$

and then multiplying both sides of the incentive constraints by the conditional probability $\pi(a|\theta)$. Thus, the linearized incentive constraints are

$$\times_{c,q} \pi(c, q, a|\theta)U(c, \theta, a) \geq \times_{c,q} \pi(c, q, a|\theta) \frac{p(q|\hat{a})}{p(q|a)} U(c, \theta, \hat{a}), \quad \forall \theta, \quad \forall a, \hat{a} \in A. \quad (4)$$

Since the choice variables are now probability distributions over $C \times Q \times A$, for each $\theta \in \Theta$, the probabilities must be restricted so that the technology function $p(q|a)$ is not violated.

The constraints

$$\forall \bar{q}, \bar{a}, \bar{\theta}, \quad \times_c \pi(c, \bar{q}, \bar{a}|\bar{\theta}) = p(\bar{q}|\bar{a}) \times_{c,q} \pi(c, q, \bar{a}|\bar{\theta}). \quad (5)$$

ensure that the exogenous production function, $p(q|a)$, is not violated. Pareto frontiers are obtained by solving the following linear program for each feasible level of participation utility, \bar{U} .

Program P1: Public-Shock Model

$$\begin{aligned} & \max_{\pi(c,q,a|\theta) \geq 0} \times_{c,q,a,\theta} \pi(c, q, a|\theta)pr(\theta)(q - c) \\ \text{s.t.} & \text{P}_{c,q,a,\theta} \pi(c, q, a|\theta)pr(\theta)U(c, \theta, a) \geq \bar{U}, \end{aligned}$$

$$(4), \quad (5),$$

$$\forall \theta, \quad \text{P}_{c,q,a} \pi(c, q, a|\theta) = 1.$$

The last constraint guarantees that the choice variable is a probability measure.

3.2 Model P2: Communication

In this model, shock θ is observed only by the agent. This change provides a role for communication to be valuable. After receiving the shock, the agent communicates with the principal by sending a report.⁷ The report allows the principal to tailor the contract to the shock the agent receives. However, alterations of the contract are limited because the agent's report on the shock is not verifiable.

By the revelation principle the reporting space R may be replaced by Θ and the message space M may be replaced with A . The agent reports his shock and the principal recommends an action. The principal chooses two parts of the contract: a randomizer that recommends action $a \in A$ as a function of received report $\theta' \in \Theta$ with probability $\pi(a|\theta')$, and a consumption schedule $\pi(c|q, a, \theta')$. Upon receiving shock θ , the agent sends report $\theta' \in \Theta$ and after receiving the recommended action he takes an action $a \in A$. The report is a function of the shock and the action is a function of the shock, the report sent, and the recommended action.

Definition 2 An incentive-compatible allocation for Model P2 is a randomizer $\pi(a|\theta)$, a consumption schedule $\pi(c|q, a, \theta)$, and deterministic strategies $\rho^*(\theta)$ and $\delta^*(a, \rho^*(\theta), \theta)$, such that, given the randomizer and the consumption schedule, strategy $\theta = \rho^*(\theta)$ and $a = \delta^*(a, \rho^*(\theta), \theta)$ satisfy

$$\prod_{c,q,a} \pi(c|q, a, \theta) p(q|a) \pi(a|\theta) U(c, \theta, a) \geq \prod_{c,q,a} \pi(c|q, a, \theta') p(q|\delta(a)) \pi(a|\theta') U(c, \theta, \delta(a)), \quad \forall \theta, \theta' \in \Theta, \forall \delta : A \rightarrow A. \quad (6)$$

In equilibrium, the agent truthfully reports his shock and then takes the recommended action.

⁷This model is a special case of Myerson (1982). A version with production shocks but without full randomization is discussed in Townsend (1993).

Incentive constraints prevent two kinds of deviations by the agent. If he reports truthfully, that is, $\theta' = \theta$, he still must be induced to take the recommended action, just as in Model P1. Checking incentive constraints for all functions $\delta(a)$, as in equation (6), is sufficient for this purpose. Many of these incentive constraints are redundant, however, because the truth-telling and obedience contract induces the agent to take the recommended action if he reports truthfully. In this case, for each recommended action it is only necessary to consider deviations from that recommended action. These incentive constraints are identical to the incentive constraints, (4), in Model P1.

The second type of deviation prevented by the incentive constraints is the agent lying about the true value of the shock θ . If the agent were to report $\theta' \neq \theta$, he must have a strategy, $\delta(a)$, of actions to take as a function of the recommended action. Thus, if he lies his utility is the sum over recommended actions of utilities obtained by taking strategy $\delta(a)$. Truth-telling is induced by calculating the set of feasible utility levels for an agent who lies and takes any strategy $\delta(a)$, and then ensuring that the value of each element in this set is less than the utility obtained from telling the truth and taking the recommended action.

The objective function and the participation constraint are linearized exactly as in Model P1. To facilitate comparison with Program P1, we break the incentive constraints into the truth-telling constraints and the moral hazard constraints along the truth-telling path. The latter constraints are identical to (4) in Program P1. The truth-telling constraints are

$$\times_{c,q,a} \pi(c, q, a|\theta)U(c, \theta, a) \geq \times_{c,q,a} \pi(c, q, a|\theta') \frac{p(q|\delta(a))}{p(q|a)} U(c, \theta, \delta(a)), \quad \forall \theta, \theta' \neq \theta \in \Theta, \forall \delta : A \rightarrow A. \quad (7)$$

Constraints which ensure that the production function are not violated and that the choice variable is a probability measure are the same as in the public-shock model.

The Pareto frontier for the communication model is calculated by solving the following linear program for all feasible values of reservation utility level, \bar{U} .

Program P2: Communication Model

$$\begin{aligned}
& \max_{\pi(c,q,a|\theta) \geq 0} \quad \times \quad \pi(c, q, a|\theta) pr(\theta)(q - c) \\
\text{s.t.} \quad & \text{P}_{c,q,a,\theta} \pi(c, q, a|\theta) pr(\theta) U(c, \theta, a) \geq \bar{U}, \\
& (4), \quad (7), \\
& \forall \bar{q}, \bar{a}, \bar{\theta}, \quad \text{P}_c \pi(c, \bar{q}, \bar{a}|\bar{\theta}) = p(\bar{q}|\bar{a}) \text{P}_{c,q} \pi(c, q, \bar{a}|\bar{\theta}), \\
& \forall \theta, \quad \text{P}_{c,q,a} \pi(c, q, a|\theta) = 1.
\end{aligned}$$

This program is identical to Program P1, except for the addition of the truth-telling constraints (7).

3.2.1 An alternative formulation

If there are more than a few points in the action grid, Program P2 is an extremely large linear program. The problem comes from the need to check all of the off-equilibrium path strategies for the truth-telling constraints. In particular, for each (θ, θ') pair the number of functions $\delta : A \rightarrow A$ is $(n_A)^{(n_A)}$, where n_A is the number of grid points in A . Essentially, the size of Program P2 grows exponentially with the number of action grid points. This severely limits the size of examples that can be solved on a computer. Fortunately, Program P2 can be reformulated into a linear program with many fewer constraints. The idea is to use interim utility to efficiently account for the off-equilibrium path strategies.

Let $\delta^*(a, \theta', \theta)$ be the optimal strategy for an agent who receives shock θ , reports $\theta' \neq \theta$, and is recommended action a . Define $w(a, \theta', \theta)$ as the utility the agent receives along this off-equilibrium path weighted by the probability of a being recommended given that θ' was

reported, that is, weighted by $\pi(a|\theta')$. Formally, $\forall a, \forall \theta, \theta' \neq \theta$,

$$w(a, \theta', \theta) = \times_{c,q} \pi(c, q, a|\theta') \frac{p(q|\delta^*(a, \theta', \theta))}{p(q|a)} U(c, \theta, \delta^*(a, \theta', \theta)).$$

Since $\delta^*(a, \theta', \theta)$ is optimal along the (a, θ', θ) path, it dominates all other possible actions.

Therefore,

$$\times_{c,q} \pi(c, q, a|\theta') \frac{p(q|\hat{a})}{p(q|a)} U(c, \theta, \hat{a}) \leq w(a, \theta', \theta), \quad \forall \theta, \theta' \neq \theta, \quad \forall a, \hat{a}. \quad (8)$$

The function $w(a, \theta', \theta)$ is an upper bound on the weighted utility a lying agent can receive along the (a, θ', θ) off-equilibrium path. Summing over these paths then gives the most utility an agent can receive from lying and choosing any strategy $\delta(a)$, as in (7). Therefore, the truth-telling constraint can be written

$$\times_{c,q,a} \pi(c, q, a|\theta) U(c, \theta, a) \geq \times_a w(a, \theta', \theta), \quad \forall \theta, \theta' \neq \theta. \quad (9)$$

The only remaining modification to Program P2 that is needed is to make $w(a, \theta', \theta)$ a choice variable.

Program P2a: Communication Model

$$\begin{aligned} & \max_{\pi(c,q,a|\theta) \geq 0, w(a,\theta',\theta)} \times_{c,q,a,\theta} \pi(c, q, a|\theta) pr(\theta)(q - c) \\ \text{s.t.} \quad & \text{P}_{c,q,a,\theta} \pi(c, q, a|\theta) pr(\theta) U(c, \theta, a) \geq \bar{U}, \\ & (4), (8), (9), \\ & \forall \bar{q}, \bar{a}, \bar{\theta}, \text{P}_c \pi(c, \bar{q}, \bar{a}|\bar{\theta}) = p(\bar{q}|\bar{a}) \text{P}_{c,q} \pi(c, q, \bar{a}|\bar{\theta}), \\ & \forall \theta, \text{P}_{c,q,a} \pi(c, q, a|\theta) = 1. \end{aligned}$$

Program P2a differs from Program P2 in that (7) has been replaced by (8) and (9), and the weighted off-equilibrium utilities have been added as a choice variable. Program P2a is also a linear program but with much fewer constraints than Program P2. The number

of truth-telling constraints in Program P2 is $n_\Theta(n_\Theta - 1)(n_A)^{n_A}$, where n_Θ is the number of elements in Θ . By contrast, the number of off-equilibrium constraints and truth-telling constraints in Program P2a is $n_\Theta(n_\Theta - 1)(n_A^2 + 1)$. There are more variables in Program P2a than in Program P2 but this cost is dwarfed by the reduction in the number of constraints, particularly for large action grids.

3.3 Model P3: No Communication

In this model, the shock and action are private information just like in Model P2. Now, however, the agent is not allowed to send a report on his shock to the principal, though the principal may still send messages to the agent. This means that all types must face the same compensation schedule since there is no way to differentiate themselves. In a deterministic version of the model, this assumption would mean that the compensation schedules are restricted to the form $c(q)$ rather than the more general form of $c(q, \theta)$ like in Models P1 and P2. A deterministic version of this model is analyzed in Holmström (1979).

For the previous two models the revelation principle justified the use of the message space A . Letting $A = M$ is no longer desirable. Instead, it is desirable to use a message space which recommends an action to the agent as a function of his shock θ . Consequently, the message space is the cross-product of state-dependent actions, $\prod_{\theta \in \Theta} A$. The Appendix contains a proof that this message space is sufficient. (Also, see Forges (1986).) An element of the new message space is a vector of state-dependent actions, \mathbf{a}_Θ . Call any such vector a strategy for the agent. Let a_θ denote the action taken if \mathbf{a}_Θ is the strategy and θ is realized.

Definition 3 *An incentive-compatible allocation for Model P3 is a randomizer $\pi(\mathbf{a}_\Theta)$, a consumption schedule $\pi(c|q, \mathbf{a}_\Theta)$, and a deterministic strategy $a_\theta = a^*(\mathbf{a}_\Theta, \theta)$ such that given the randomizer, the consumption schedule, and the recommended strategy \mathbf{a}_Θ , the strategy*

$a_\theta = a^*(\mathbf{a}_\Theta, \theta)$ satisfies

$$\prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) p(q|a_\theta) U(c, \theta, a_\theta) \geq \prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) p(q|\hat{a}) U(c, \theta, \hat{a}), \quad \forall \hat{a} \in A, \quad (10)$$

for each realization of θ .

Writing the principal-agent problem as a linear program is difficult because of the cumbersome message space. Frequent use is made of the identity

$$\pi(c, q|\mathbf{a}_\Theta) = \pi(c|q, \mathbf{a}_\Theta) \prod_{\theta} \overset{\bar{A}}{\text{pr}}(\theta) p(q|a_\theta) \quad ! \quad (11)$$

The term $\prod_{\theta} \overset{\bar{P}}{\text{pr}}(\theta) p(q|a_\theta)$ is the probability of output q , before θ is realized, given strategy \mathbf{a}_Θ , that is, $\text{prob}(q|\mathbf{a}_\Theta)$. The unconditional probability of consumption c , output q , and strategy \mathbf{a}_Θ , is $\pi(c, q, \mathbf{a}_\Theta) = \pi(c, q|\mathbf{a}_\Theta) \pi(\mathbf{a}_\Theta)$.

The principal's utility is

$$\prod_{\mathbf{a}_\Theta} \prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) \prod_{\theta} \overset{\bar{A}}{\text{pr}}(\theta) p(q|a_\theta) (q - c) \pi(\mathbf{a}_\Theta),$$

while the agent's utility is

$$\prod_{\mathbf{a}_\Theta} \prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) \prod_{\theta} \overset{\bar{A}}{\text{pr}}(\theta) p(q|a_\theta) U(c, \theta, a_\theta) \pi(\mathbf{a}_\Theta).$$

Using equation (11), the principal's utility function simplifies to

$$\prod_{c,q,\mathbf{a}_\Theta} \pi(c, q, \mathbf{a}_\Theta) (q - c),$$

and the agent's utility function simplifies to

$$\prod_{c,q,\mathbf{a}_\Theta} \pi(c, q, \mathbf{a}_\Theta) \prod_{\theta} \frac{\overset{\bar{P}}{\text{pr}}(\theta) p(q|a_\theta)}{[\overset{\bar{P}}{\text{pr}}(\theta) \text{pr}(\theta) p(q|a_\theta)]} U(c, \theta, a_\theta).$$

Both equations are linear in $\pi(c, q, \mathbf{a}_\Theta)$. The ratio $\frac{\overset{\bar{P}}{\text{pr}}(\theta) p(q|a_\theta)}{[\overset{\bar{P}}{\text{pr}}(\theta) \text{pr}(\theta) p(q|a_\theta)]}$ in the agent's utility function is the likelihood that given recommended strategy \mathbf{a}_Θ and output q , the agent

took action a_θ .⁸

Incentive constraints are similarly linearized. Using $p(q|\mathbf{a}_\Theta) = \prod_\theta pr(\theta)p(q|\mathbf{a}_\Theta)$, first substitute

$$\pi(c|q, \mathbf{a}_\Theta)p(q|a_\theta) = \pi(c, q|\mathbf{a}_\Theta) \frac{p(q|a_\theta)}{\prod_\theta pr(\theta)p(q|a_\theta)}$$

into equation (10). Multiply both sides of each incentive constraint by $\pi(\mathbf{a}_\Theta)$, the probability of the principal recommending strategy \mathbf{a}_Θ , to obtain

$$\begin{aligned} \forall \mathbf{a}_\Theta, \forall \theta, \quad & \times_{c,q} \pi(c, q, \mathbf{a}_\Theta) \frac{p(q|a_\theta)}{\prod_\theta pr(\theta)p(q|a_\theta)} U(c, \theta, a_\theta) \\ & \geq \times_{c,q} \pi(c, q, \mathbf{a}_\Theta) \frac{p(q|\hat{a})}{\prod_\theta pr(\theta)p(q|a_\theta)} U(c, \theta, \hat{a}) \quad , \forall \hat{a} \in A. \end{aligned} \quad (12)$$

The ratios $\frac{p(q|a_\theta)}{\prod_\theta pr(\theta)p(q|a_\theta)}$ and $\frac{p(q|\hat{a})}{\prod_\theta pr(\theta)p(q|a_\theta)}$ are similar to the likelihood ratio $\frac{p(q|\hat{a})}{p(q|a)}$ in Models P1 and P2, except now when the principal observes output q he makes inferences without knowing the realization of θ .

Next, the technology constraints are incorporated. Consistency of allocations with the exogenous technology $p(q|a)$ is accomplished by

$$\forall \bar{q}, \bar{\mathbf{a}}_\Theta, \quad \times_c \pi(c, \bar{q}, \bar{\mathbf{a}}_\Theta) = \left[\prod_\theta pr(\theta)p(\bar{q}|\bar{\mathbf{a}}_\Theta) \right] \times_{c,q} \pi(c, q, \bar{\mathbf{a}}_\Theta). \quad (13)$$

The objective function and each set of constraints are now linear. The Pareto frontier for the no-communication model is calculated by solving the following linear program for all feasible values of reservation utility, \bar{U} .

⁸If preferences are of the form $U(c) + V(\theta, a)$, as they will be in later numerical examples, the agent's utility function simplifies to

$$\times_{c,q,\mathbf{a}_\Theta} \pi(c, q, \mathbf{a}_\Theta) [U(c) + \prod_\theta pr(\theta)V(\theta, a_\theta)].$$

Program P3: No-Communication Model

$$\begin{aligned}
& \max_{\pi(c, q, \mathbf{a}_\Theta) \geq 0} \sum_{c, q, \mathbf{a}_\Theta} \pi(c, q, \mathbf{a}_\Theta)(q - c) \\
\text{s.t. } & \sum_{c, q, \mathbf{a}_\Theta} \pi(c, q, \mathbf{a}_\Theta) \sum_{\theta} \frac{p(\theta)p(q|a_\theta)}{p(\theta)p(q|a_\theta)} U(c, \theta, a_\theta) \geq \bar{U}, \\
& (12), (13), \\
& \sum_{c, q, \mathbf{a}_\Theta} \pi(c, q, \mathbf{a}_\Theta) = 1.
\end{aligned}$$

The last constraint in the program ensure that the choice variable is a probability measure.

4 Results

Optimal contracts are calculated by solving Programs P1, P2a, and P3. Given a feasible level of the agent's utility, \bar{U} , communication's value is the difference between the values of the objective functions for the communication and no-communication models. A solution to the communication model weakly dominates a solution to the no-communication model because no-communication contracts can always be replicated with a communication contract. Similarly, a solution to the public-shock model weakly dominates that of the communication model because Program P2 only differs from Program P1 in the addition of the truth-telling constraints. Of more interest are communication's effects on the optimal contract and conditions which make communication strictly valuable. Examples are provided which demonstrate how consumption risk sharing and labor supply are affected by communication. For the finite-action case, a sufficient condition for communication to be valuable is provided.

Computer memory limits the size of linear programs that can be solved. It was for this reason Program 2a was developed. Nevertheless, reasonably large programs are computable. To keep the linear programs to a manageable size the set of types, Θ , was limited to two elements. Other sets were not so limited. Consumption grids were set to have at least

100 elements in order to generate a reasonable amount of curvature. Action and output grids were set up to ten grid points in some examples.⁹ The programs were programmed into GAMS, a modeling language, which called Minos, a non-linear program solver which also solves linear programs. The programs were solved on an IBM Risc6000 workstation with 256 MB of RAM.

Most of the analysis assumes that the principal is risk neutral and that the agent is risk averse with a utility function separable in consumption and effort, and in which the marginal disutility of effort is increasing with θ . Proposition 3, contained in the Appendix, proves that with these preferences a large number of strategies are not incentive compatible in the no-communication model, greatly reducing the grid size that needs to be considered when solving Program P3.

In order to develop intuition, a simple two action case is studied first. In this case, the bindingness of incentive constraints in the no-communication model are important in determining whether communication is valuable. If communication is valuable, it is valuable because it improves consumption risk sharing. This result is generalized to the case of any finite number of actions. Next, an example with many more actions and outputs is studied. Communication is valuable for similar reasons, even in cases when the no-communication bindingness condition on incentive constraints does not hold. This example also demonstrates how communication can alter labor supply, lowering it relative to the no-communication example. Finally, two additional examples are studied that illustrate how lotteries may help to separate agent types. The latter of these two examples uses non-standard preferences, thus illustrating that the computational methods are applicable to arbitrary specifications of preferences and technology.

⁹Computer hardware memory limits were not reached in computing these examples. If these memory limits are reached decomposition methods can be used at some cost in speed. See Prescott (2000) for an application of these methods to solving moral-hazard linear programs.

4.1 Two-Action Case

For the two-action case there are essentially two types of equilibria, which are differentiated by the actions chosen by the agents.

Definition 4 *A pooling equilibrium is when both types of agents take the same action. A separating equilibrium is when the two types of agents take different actions.*

This definition excludes lotteries over assigned actions. The lottery case is considered a linear combination of the defined kinds of equilibria.

All three models were solved numerically for the following choice of parameters: $A = [0, 5]$, $\Theta = [5, 10]$, $Q = [2, 20]$, and C gridded at intervals of 0.1 over the range $[0, 20]$. Let the subscript $i = l, h$ index low and high actions, outputs, and shocks. The utility function is $U(c, \theta, a) = c^{0.5} + (\theta - a)^{0.5}$. The probability distribution of output given the action is

	q_l	q_h
a_l	0.8	0.2
a_h	0.2	0.8

The probabilities of the shocks are $pr(\theta_l) = .25$, $pr(\theta_h) = .75$.

Figures 2 and 3 describe optimal consumption schedules for Models P1 and P2 as a function of the output, the shock, and the reservation utility level \bar{U} . The vertical axis is expected consumption. Expected consumption is reported because with finite numbers of consumption grid points, solutions are often lotteries over adjacent grid points. The range of ex-ante utilities shown are those which cause a separating equilibrium in which the low- θ agent works a_l and the high- θ agent work a_h for both models. Figure 2 describes consumption as a function of output for the low- θ (high marginal disutility of effort) agent. The two top lines, labeled Model P1, describe consumption in the public-shock model as a function of output and \bar{U} . These two lines are identical so for each \bar{U} the agent's consumption is fully insured over the output. The bottom two lines, labeled Model P2, describe consumption in the communication model as a function of output and \bar{U} . Again,

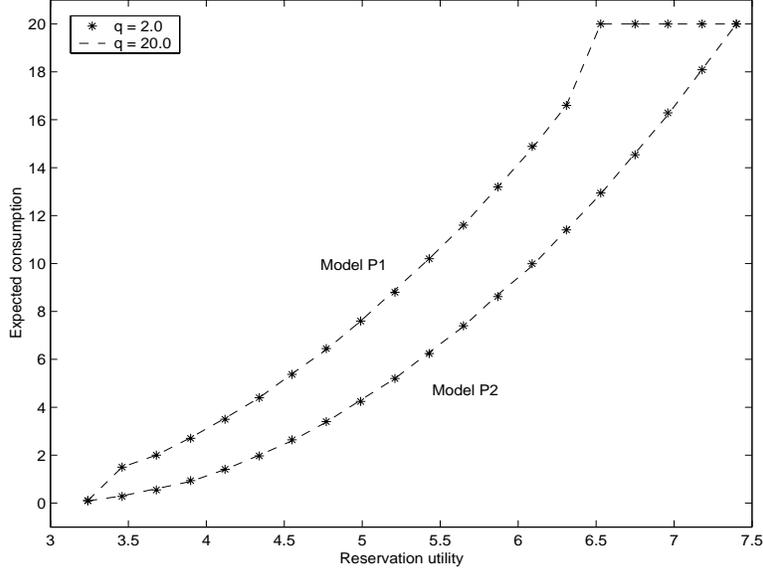


Figure 2: Two Action Case – Contracts for low- θ agents in Models P1 and P2 as a function of reservation utility level \bar{U} .

the two lines are identical and agents are fully insured over output. In both models, agents are fully insured over output because they are taking the low action and there are no binding incentive constraints at this action level. The only difference between the models is the level of consumption.

Figure 3 describes consumption schedules for Models P1 and P2 when the agent receive the high shock and is recommended the high action.¹⁰ When the high- θ agent is recommended the high action he must be induced to work hard. Consequently, if the agent produces the high output he is rewarded high consumption. Dashed lines describe the communication contracts and dotted lines describes the public-shock contracts as they vary with reservation utility level \bar{U} . The two top lines are compensation if the high output occurs and the two bottom lines are compensation if the low output occurs. Again, the only difference between the two models is the level of consumption.

¹⁰For portions of the reported range of \bar{U} , the high shock agent was recommended a lottery over the high and low efforts. What is shown in the figure is the contract if the high-shock agent is recommended the high effort.

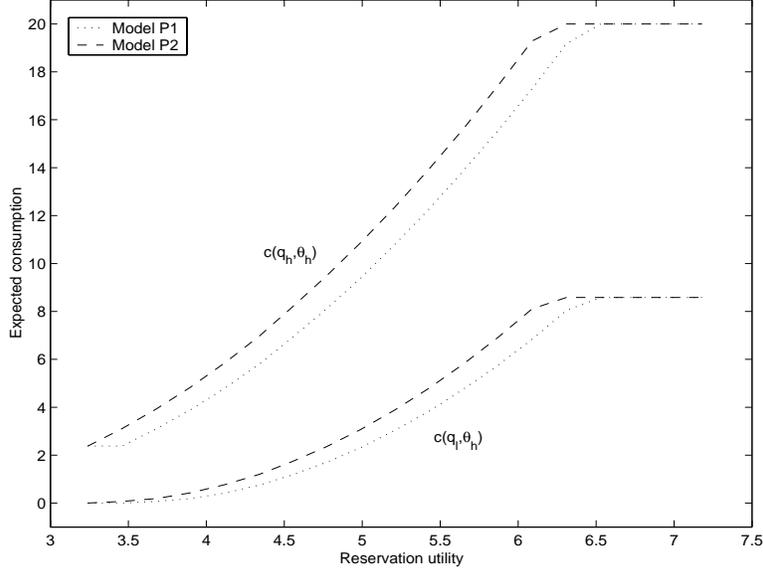


Figure 3: Two Action Case – Contracts for high- θ agents in Models P1 and P2 as a function of reservation utility level \bar{U} .

Compared with Model P1, Model P2 lowers consumption in the bad state while it raises it in the good state. The greater spread over the states is a direct effect of the hidden information. The principal prefers the Model P1 contract, but if he offers it the high- θ agent would lie and take the low action. To induce truth-telling, consumption is lowered in the bad state and raised in the good state. Consequently, the solution to the communication model does not insure the agent over the states as well as the solution to the public-shock model.

The no-communication contract is not explicitly shown. However, it is identical to the contract the high- θ agent faces in the communication model. Agents of both types face the same contract because, unlike with communication, the principal cannot tailor the compensation schedule to the agent's type. The gain from communication in this example is the value to the principal of providing the same expected utility to the low- θ agent with constant consumption rather than with the more variable consumption in the

no-communication contract.

This logic is not just specific to this example.

Proposition 1 *With two actions, if there is a separating equilibrium for the no-communication model then communication is valuable.*

Proof: For clarity assume that the consumption set is a continuum. Since preferences are separable, utility is concave in consumption, and the shock θ does not affect utility from consumption, there will not be a consumption lottery, and compensation can be written $c(q)$. In a separating equilibrium for the no-communication model, P3, the two incentive constraints take the following form,

$$\times \quad p(q|a_h)U(c(q)) + V(\theta_h, a_h) = \times \quad p(q|a_l)U(c(q)) + V(\theta_h, a_l), \quad (14)$$

$$\times \quad p(q|a_l)U(c(q)) + V(\theta_l, a_l) > \times \quad p(q|a_h)U(c(q)) + V(\theta_l, a_h). \quad (15)$$

Only the incentive constraint for the high- θ agent binds, the other constraint does not hold at equality because the low- θ agent has a higher marginal disutility of labor.

Construction of a Pareto-superior communication contract proceeds by showing that there exists a communication contract which leaves agent's interim utilities unchanged (and consequently, leaves ex-ante utility \bar{U} unchanged) but which improves the principal's utility. Index the contract by the state. Write it as $c(q, \theta)$. Let $c(q, \theta_h) = c(q)$, and let $c(q, \theta_l) = \hat{c}$ such that

$$U(\hat{c}) = \times \quad p(q|a_l)U(c(q)). \quad (16)$$

This contract gives the low- θ agent a constant transfer and does not change the utility either type of agent receives. This contract is also incentive compatible. If the high- θ agent tells the truth he will not take the low action because his incentives are unchanged from the Model P3 contract. Substituting equation (16) into equation (14) shows that this agent

will not lie and then take the low action. He will also not lie and take the high action because this is even worse than lying and taking the low action. Reporting the truth and then taking the high action is not beneficial for the low- θ agent because with a constant transfer there is no gain to working hard. He will also not lie because equations (15) and (16) imply,

$$\begin{aligned}
\underset{q}{\times} p(q|a_l)U(\hat{c}) + V(\theta_l, a_l) &= \underset{q}{\times} p(q|a_l)U(c(q)) + V(\theta_l, a_l) \\
&= \underset{q}{\times} p(q|a_l)U(c(q, \theta_h)) + V(\theta_l, a_l) \\
&> \underset{q}{\times} p(q|a_h)U(c(q, \theta_h)) + V(\theta_l, a_h), \tag{17}
\end{aligned}$$

which means that both of his off-equilibrium path incentive constraints are satisfied. This feasible contract is better for the principal because concavity of the utility function and equation (16) imply that $\hat{c} < \mathbf{P}_q p(q|a_l)c(q)$. Q.E.D.

No general statement can be made on the value of communication when the solution to the no-communication game is a pooling equilibrium. Whether the value of communication is greater than zero for pooling equilibrium depends on the parameters. The author has generated examples, where for a portion of the range of \bar{U} that produce a high-action pooling solution to the no-communication model, communication is valuable. The optimal contract with communication recommends low effort to the low- θ agent with positive probability. Along the rest of the range, communication is not valuable. Similar counter-examples have been found for low-action pooling equilibrium.

Communication in the example did not change expected output nor did it change the actions taken; it only allocated consumption more efficiently over states. In the example shown the gain from communication as a percentage of expected output is not very much, about 0.5 percent. However, by raising the probability of the lower state the gain to communication can be raised to at least 5 percent of expected output.

4.2 A generalization

There is a gain to communication in the two-action example because communication allows the principal to take advantage of the lack of an incentive constraint on the low agent. This result can be extended to cases where the low- θ agent is not taking the lowest action.

The next proposition is proven for preferences of the form $U(c, \theta, a) = U(c) + V(\theta, a)$, with U strictly concave, $V_a < 0$, $V_{aa} < 0$, and $V_{\theta a} > 0$.

Proposition 2 *Assume that the preferences above hold, that there are two outputs with the probability of the high output increasing with the action, and that in the solution to the no-communication model the incentive constraint does not bind on the low- θ agent. Then, communication is valuable.*

Proof: See the Appendix.

The proof is very similar to that of the two-action case. A Pareto-superior communication contract is constructed by altering the compensation of the low- θ agent to improve insurance without affecting incentives. The difference is that the low- θ agent is only insured to the extent allowed by his incentive constraints.

4.3 Many-Action Case

In the previous examples communication is valuable because it allows the principal to alter the contract along the θ_l path without affecting the incentive constraints. The proofs relied on the recommended actions not changing when communication was added. With a continuum, or even a large, finite, number of actions, communication should also change labor allocations.

For cases with indivisibilities, or possibly severe non-convexities in production, the two-action case may be a reasonable abstraction and the results will be applicable. However, if the desired application demands more actions then not only will labor supply change but the

previous constructions of Pareto-superior communication contracts no longer can be used. For example, in the no-communication model with a continuum of actions, continuum of consumption points, and technology $p(q_h|a)$ concave in a , then the action incentive constraints can be replaced by the first-order condition

$$\times \frac{\delta(p(q|a_\theta))}{\delta a} U(c(q)) = V'(\theta - a_\theta), \text{ if } a_\theta \in (0, \bar{A}), \forall \theta, \quad (18)$$

where $(0, \bar{A})$ is the interior of set A .¹¹

The next example demonstrates the effects of adding communication to labor supply. Both consumption schedules and labor allocations change. This example uses the following parameter values: $C = [0.00, 0.02, 0.04, \dots, 2.00]$, $A = [0.0, 0.08, 0.16, \dots, 0.80]$, $Q = [-2.0, -1.6, -1.2, \dots, 2.0]$, and $\Theta = [2.0, 0.5]$. The probability of the shocks are $pr(\theta = 2.0) = .25$, and $pr(\theta = 0.5) = .75$. The technology is roughly a normal distribution where effort controls the mean. In particular, $p(q|a) = (1/\sqrt{2 * 3.1416})e^{-0.5*(q-1.75*a)^2} / (\sum_q p(q|a))$. The denominator normalizes the probabilities to sum to one. Utility is $U(c, a, \theta) = c^{0.5} - \theta a^2$. Unlike in the previous example, a low- θ value corresponds to a low marginal disutility of effort.

Programs P1, P2a, and P3 were solves for $\bar{U} = 0.3578$. In all three models, optimal action choices were degenerate for both types of agents. In Models P1 and P2, $a(\theta = 2.0) =$

¹¹The first-order approach, that is representing the incentive constraints with a first-order condition, is valid if there are only two outputs. If there are more than two outputs this need not be true, even with the standard assumptions made on the technology. Proofs of the validity of this substitution in the moral-hazard problem usually rely on showing that consumption schedules are increasing in output, which can be guaranteed by assuming that the technology satisfies a monotone likelihood ratio property. Monotonicity, combined with a convexity condition on the distribution function, ensures that the agent's local optimum is also the global optimum. See, for example, Hart and Holmström (1987). In this model if there are more than two outputs, the assumptions on technology are not sufficient. The first-order necessary condition on consumption is

$$\frac{1}{U^0(c(q_j))} = \lambda + \frac{\sum_{\theta} \mu_{\theta} p^0(q_j|a_{\theta})}{\sum_{\theta} pr(\theta) p(q_j|a_{\theta})},$$

where λ is the multiplier on the participation constraint, μ_{θ} are the multipliers on agent- θ 's incentive constraint, and a_{θ} is agent- θ 's action. The monotone likelihood ratio property does not guarantee a monotonic consumption sharing rule (This is also true for the communication model, P2.)

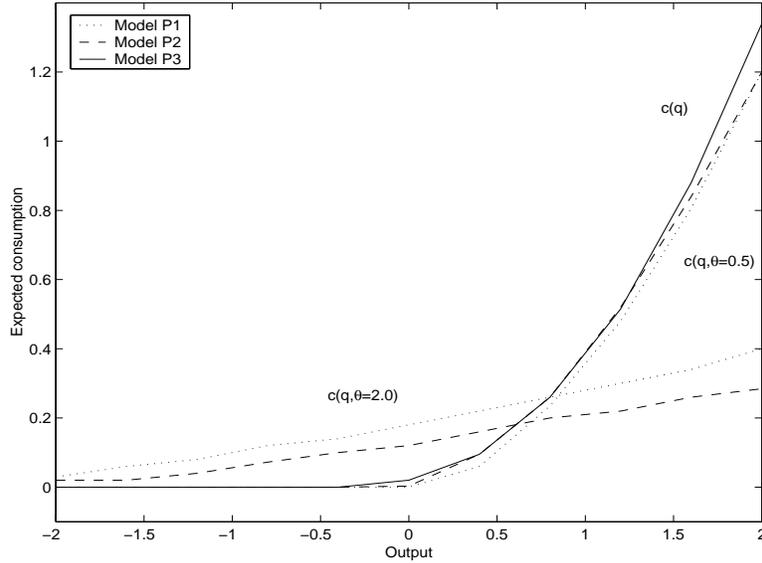


Figure 4: Many-Action Case – Expected consumption as a function of the state and the output for $\bar{U} = 0.3578$.

0.08 and $a(\theta = 0.5) = 0.56$. For Model P3, the high marginal disutility of effort agent, $\theta = 2.0$, works more than in the other models. His effort is $a(\theta = 2.0) = 0.16$. For the other shock, the solution is $a(\theta = 0.5) = 0.56$.

Optimal consumption schedules for all three models are reported in Figure 4. The solid line is the consumption schedule for the no-communication model. Of course, both types of agents face this schedule. The dashed lines are the consumption schedules for the communication model. The curved line that closely follows the no-communication schedule is the schedule for the low marginal disutility of effort agent, that is, $\theta = 0.5$. It provides less consumption than the no-communication schedule. The nearly linear dashed line is the consumption schedule for high marginal disutility of effort agent, that is, $\theta = 2.0$. Comparing the no-communication and communication models, the no-communication model does not provide as much consumption insurance to the agent. It also requires the $\theta = 2.0$ agent to work harder. For these reasons, it provides a higher level of consumption.

The effect of the truth-telling constraints can be seen by comparison of the communication contracts with the public-shock contracts. Actions are the same for the two models so a comparison can focus solely on the consumption schedules. The dotted line describes the optimal public shock consumption schedules. Qualitatively, it is similar to those of the communication model. However, the quantitative differences illustrate the effect of the truth-telling constraints in Model P2. The public-shock model solution strictly dominates in welfare terms the communication model in this example so Program 2a would like to implement that allocation. It is not, however, incentive compatible. Like in the two-action example, the consumption level provided is high enough that the $\theta = 0.5$ agent would like to report that he experienced the bad shock and then work some low amount. To prevent this, Program 2a lowers the level of compensation paid to an agent who receives the bad shock and it raises the level of compensation paid to the agent who receives the good shock. The truth-telling constraints limit insurance across the shocks.

Despite the program's ability to insure the agent across the shocks, communication still allows some insurance, at least, relative to the no-communication model. Figure 5 reports the agent's utility conditional on the shock he receives for the full range of feasible \bar{U} levels, that is, for the entire Pareto frontier. The solid lines report the conditional utility for the no-communication model while the dashed lines report the conditional utility for the communication model. For both models, the lower line corresponds to $\theta = 2.0$ and the higher line corresponds to $\theta = 0.5$. There is a bigger spread between the utilities for the no-communication model than for communication model.¹²

Communication also effects labor supply for much of the range of reservation utilities \bar{U} , not just for the specific \bar{U} solution reported above. Figure 6 reports expected action conditional on the shock over the Pareto frontier. Again, solid lines correspond to the

¹²At the highest level of \bar{U} , there is no spread in either model because all types are assigned an action level of $a = 0.0$ and since effort level is multiplicative with the shock θ , both types receive the same utility.

no-communication model, and dashed lines correspond to the communication model. The public-shock model is illustrated with stars connected by a dotted line. The lower set of lines are optimal expected actions conditional on receiving the bad shock and the higher set of lines correspond to the good shock. Comparing the public-shock model with the communication model, it is apparent that there is decline in effort levels. Because the communication model cannot provide as much insurance as the public shock model, consumption schedules are altered and the $\theta = 0.5$ agent generally works a little less than in the public shock model. There are also labor supply effects for the no-communication model. In this example, effort in the no-communication model is always weakly less than in the communication model for $\theta = 0.5$ agents and it is weakly more for $\theta = 2.0$ agents. As suggested by Figure 4, Program 3 is tailoring compensation schedules towards the low-marginal disutility of effort agent. This raises effort of the other type for much of the Pareto frontier. Similarly, to ensure that the agent receives \bar{U} utils, the program is also being more generous in consumption and allowing slightly less effort for the low-marginal disutility of effort agent.¹³ Finally, for certain values of \bar{U} the incentive constraint on the low-marginal disutility of effort agent binds in the solution to the no-communication model, yet communication is still valuable.

4.4 Lottery examples

The following two examples demonstrate how lotteries can mitigate truth-telling constraints. In the first example a lottery over recommended actions is part of the optimal communication contract and in the second example a consumption lottery is part of the optimal communication contracts.

The action-lottery example uses the following parameter values: $C = [0.0, 0.2, 0.4, \dots, 20.0]$,

¹³This property is not a theorem. There are examples where effort of the bad shock agent is less in the no-communication model than in the communication model.

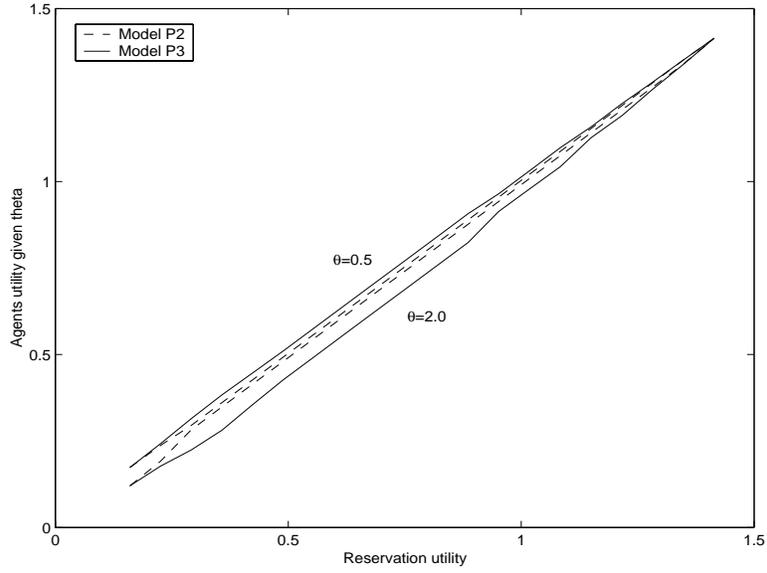


Figure 5: Many-Actions Case – Expected action as a function of the state and the reservation utility level \bar{U} .

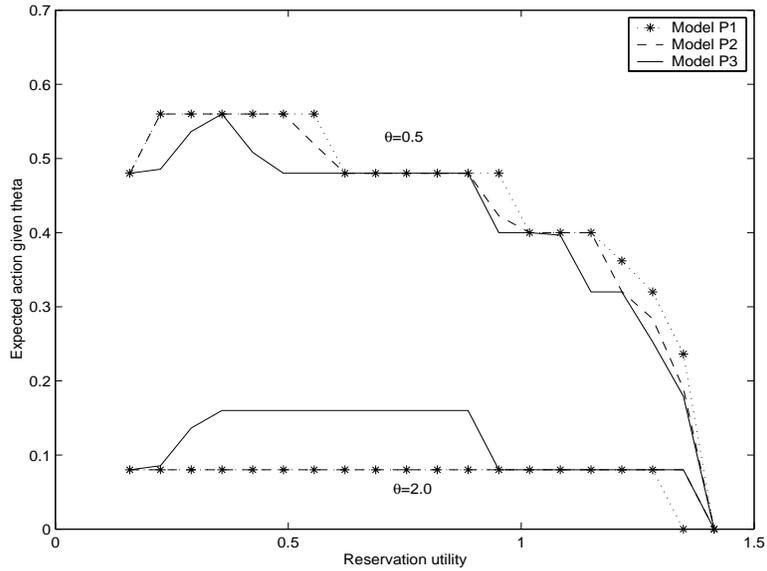


Figure 6: Many-Actions Case – Agent’s expected utility as a function of the state and the reservation utility level \bar{U} .

$A = [1.50, 1.85, 2.20, \dots, 5.00]$, two output levels $Q = [2.0, 20.0]$, and $\Theta = [5.0, 10.0]$. The probability of the shocks are $pr(\theta = 2.0) = .25$, and $pr(\theta = 10.0) = .75$. The probability of the high output is $p(q = 20.0|a) = (a/6)^{0.8}$, which means that the probability of a high output is increasing and concave in the action. Utility is $U(c, a, \theta) = c^{0.5} + (\theta - a)^{0.5}$, so as in the two-action example a high θ means a low marginal disutility of effort.

Programs P1, P2a, and P3 were solved for $\bar{U} = 5.0$. For the public-shock model, the solution is reported in Table 2. As indicated by the fourth and fifth columns, action assignments are degenerate in agent types. The $\theta = 5.0$ agent (high marginal disutility of effort) is assigned $a = 2.90$ with probability one and the $\theta = 10.0$ agent is assigned the higher action $a = 5.0$ with probability one. Optimal contracts along the $\theta = 5.0$ path are characterized by small compensation for $q = 2.0$ (the low output) and much higher compensation for the $q = 20.0$ output.¹⁴ The contract for the $\theta = 10.0$ agent is similar though his level of consumption is lower than that of the other agent type.

$\pi(c, q, a \theta)$	c	q	a	θ
0.1128	3.00	2.0	2.90	5.0
0.3283	3.20	2.0	2.90	5.0
0.5590	15.20	20.0	2.90	5.0
0.0349	2.40	2.0	5.00	10.0
0.1009	2.60	2.0	5.00	10.0
0.5291	10.00	20.0	5.00	10.0
0.3352	10.20	20.0	5.00	10.0

Table 2: Action-lottery case – Solution to Program P1, the public-shock model.

The solution to Program P1 is not incentive compatible in the communication model. The $\theta = 10.0$ type agent would report he was the other type and take the better compensation schedule. The solution to Program 2a is reported in Table 3.

¹⁴There is a consumption lottery in the compensation schedule for this type of agent if he produces the low output. This lottery, however, is only an artifact of the consumption grid. With a fine enough grid it would disappear. In the next example, there will be a substantive consumption lottery.

$\pi(c, q, a \theta)$	c	q	a	θ
0.3889	5.80	2.0	1.50	5.0
0.1914	5.80	20.0	1.50	5.0
0.1468	1.40	2.0	2.90	5.0
0.0383	1.60	2.0	2.90	5.0
0.1729	11.00	20.0	2.90	5.0
0.0617	11.20	20.0	2.90	5.0
0.0694	2.80	2.0	5.00	10.0
0.0663	3.00	2.0	5.00	10.0
0.1705	10.60	20.0	5.00	10.0
0.6938	10.80	20.0	5.00	10.0

Table 3: Action-lottery case – Solution to Program P2a, the communication model.

To preserve incentive compatibility Program 2a does several things. First, it slightly raises compensation to the $\theta = 10.0$ type. This increase makes truthful reporting more attractive. Second, it lowers compensation to the low type and it lowers his labor effort by recommending the lowest action a positive fraction of the time. This is a non-trivial action lottery. There are several actions between 1.50 and 2.90 that the program could have chosen but did not. One reason for this lottery is a non-convexity created by the moral hazard constraint. Actions near 1.50 are either too expensive to implement or infeasible so the program switches to the 1.50 action. However, the 2.55 action is implementable and the public shock model, which also has the moral hazard non-convexity implements this action for a range of the Pareto frontier. Consequently, because the communication model chooses an action lottery over implementing this action, the lottery appears to help with the truth-telling constraints. The very low action is a way of transferring utility to the low- θ agent in a way that is relatively less appealing to the high- θ agent.

The final example demonstrates that consumption lotteries may also mitigate truth-telling constraints. In this example, the preference shock is a shock to risk aversion. Consumption lotteries will then be valuable to separate agents on this basis. The example is similar to Cole (1989) but with the addition of a moral hazard problem.

This example uses the following parameter values: $C = [0.01, 0.02, 0.03, \dots, 2.00]$, $A = [0.2, 0.4, 0.6, 0.8]$, $Q = [0.0, 10.0]$, and $\Theta = [0.10, 0.99]$. The probability of the shocks are $pr(\theta = 2.0) = 0.5$, and $pr(\theta = 0.5) = 0.5$. The probability of the high output is $p(q = 2.0|a) = a^{0.8}$. Utility is $U(c, a, \theta) = (c + 0.4)^\theta / \theta - 0.6a^2$. The high- θ agent is virtually risk neutral in this example. Finally, Programs P1, P2a, and P3 were solved for a reservation level of utility of $\bar{U} = 5.4$.

$\pi(c, q, a \theta)$	c	q	a	θ
0.0719	0.00	0.0	0.80	0.10
0.0916	0.01	0.0	0.80	0.10
0.8365	0.72	10.0	0.80	0.10
0.0217	0.30	0.0	0.80	0.99
0.1417	0.40	0.0	0.80	0.99
0.2030	1.29	10.0	0.80	0.99
0.6335	1.39	10.0	0.80	0.99

Table 4: Consumption-lottery case – Solution to Program P1, the public-shock model.

The solution to Program P1, the public-shock model, is reported in Table 4. Both types of agents are induced to work the highest effort. This is done by compensating them well if they produce the high output. There are consumption lotteries for the nearly risk neutral type ($\theta = 0.99$) but these appear to be due to numerical approximation. The main difference between the two types' contracts is that the nearly risk neutral type receives a uniformly higher level of consumption. This difference is due to the effect of θ on the marginal utility of consumption. Program P1 tailors compensation to the two types in the most cost-effective way.

Because the $\theta = 0.99$ type's contract is strictly higher than that of the other type, the other type would have an incentive to misreport in the communication model. To prevent this misreporting, Program P2a resorts to a consumption lottery to separate the types. The solution to Program P2a is reported in Table 5.

$\pi(c, q, a \theta)$	c	q	a	θ
0.0222	0.06	0.0	0.80	0.10
0.1413	0.07	0.0	0.80	0.10
0.3082	0.87	10.0	0.80	0.10
0.5283	0.88	10.0	0.80	0.10
0.1635	0.00	0.0	0.80	0.99
0.2988	0.00	10.0	0.80	0.99
0.5377	2.00	10.0	0.80	0.99

Table 5: Consumption-lottery case – Solution to Program P2a, the communication model.

As in the solution to Program P1, the highest action levels are implemented for both types of agents. Now, however, the contract changes in three ways to satisfy truth-telling constraints: First, the compensation level of the risk-averse type ($\theta = 0.10$) is higher. Second, the consumption level as measured by expected consumption of the $\theta = 0.99$ agent declines. And, third, the way in which this agent is paid changes. Now, if he produces the high output $q = 20.0$ he receives a lottery over the highest and lowest possible levels of consumption. This randomization barely affects his utility but has a strong impact on the truth-telling constraints for on the other, more risk averse agent.

Finally, if communication is removed then the consumption lotteries disappear. The solution to Program P3 is reported in Table 6. Again, both agent types work the highest effort level. Now the compensation contract cannot be tailored to the agent type. Relative to the Program P1 solution, consumption levels are higher for the $\theta = 0.10$ type and lower for the $\theta = 0.99$ type.

$\pi(c, q, a_{\theta_1}, a_{\theta_2})$	c	q	a_{θ_1}	a_{θ_2}
0.1601	0.11	0.0	0.80	0.80
0.0777	0.12	0.0	0.80	0.80
0.1392	1.07	10.0	0.80	0.80
0.6973	1.09	10.0	0.80	0.80

Table 6: Consumption-lottery case – Solution to Program P3, the no-communication model.

5 Field Work

Four of the cropping group arrangements were fixed rent while four were sharecropping arrangements, with a fifty percent share. As has long been noted in the literature, principal-agent theory predicts more insurance than is found in an explicit fixed rent contracts, and only under the most restrictive conditions does it predict the linear consumption sharing rules found in an explicit sharecropping arrangement.

One finding from the surveys and the follow-up fieldwork is that the explicit terms of these contracts missed important features. For example, respondents regularly reported that fixed rents would be reduced in the event of drought or crop failure. Even in the absences of these extreme shocks there are implicit terms. Of the four groups with fixed rents, three had a dispute with their landowner over the final amount of rent to pay. These disputes were settled with the help of outsiders.¹⁵

Contingencies need not only be introduced through rental payments. In the sharecropping arrangements, two of the groups had disputes over the amount of work supplied by a tenant. Plus, there were cases where input purchases were made in a different proportion than in the original agreement. Finally, it should be noted that insurance need not only come from the landowner. Several of the groups used outside financiers and implicit contingencies in these loans can be a source of insurance and incentives.

Implicit contractual terms also seem to matter for assessing communication. The surveys indicated that, when possible, there was frequent communication from the tenants to their landowners and to their financiers. At each interview, groups were asked when they had last met with their landowner or financier to discuss crop progress. For four of the groups this is irrelevant since their landowner also participated in the group so saw the

¹⁵There are actually two different types of fixed rent contracts in Aurepalle, each with its own name. In one, tenants pay the fixed rent at the end of season. In the other, less common one, tenants pay the rent in advance. The latter type of contract appears to provide less scope for implicit insurance. The fixed rent cropping groups paid their rents at the end of the season.

group members regularly. Of the other four, three were absentee landowners who lived in the city. All three arrangements with these landowners were fixed rent contracts. Communication with these landowners is prohibitively costly because of the distances and a lack of telephone connections to Aurepalle. Only once did any of these groups communicate with their landowner.¹⁶ However, two of these groups financed their production through a financier and both indicated they met with their financier regularly. One group indicated that their meetings had occurred between four and ten days before each interview. Another group responded that their meetings had occurred six to ten days before all but one interview. At this interview they responded that no meeting had occurred. The final group, with the resident non-participating landowner, met frequently with their landowner, usually two or three days before each interview. This group did not have an outside financier.

The precise role of communication in the contracts is harder to ascertain. Certainly crop progress and sometimes choice of inputs were discussed but how they were used in assessment is not clear. That some information was transmitted is apparent along other dimensions. The surveys measure ex post information sets of landowners in order to compare them with those of the tenants. Landowners generally knew most of the information. Whether this is the result of communication as in Model P2, the information being public as in Model P1, or monitoring and inspections as in costly state verification models is unknown. Also, financiers were not interviewed as part of the surveys, though they did not inspect operations. Follow-up interviews indicate that they are in frequent contact with their borrowers.

Still, while unusual, one contractual arrangement is strikingly like the theory. This arrangement was identified during the follow-up visit by the author. This visit identified

¹⁶In the follow up interviews, a group not originally surveyed was identified that had a bad year and wanted to renegotiate the payment. They did not renegotiate because it would have been too expensive to travel to the city of Hyderabad (where the landowner lived) and discuss the matter with him. Consequently, they paid the full rent.

a cropping group with an explicit contract that resembles the contracts in Models P1 and P2. This group sharecropped land from a non-resident landowner. At harvest time this landowner would come to the village to collect his share. At this point, the tenants could ask him to plant a crop of paddy (rice) for the following dry season. The landowner would evaluate the well level (rice needs irrigation) and ask other tenants about this group. Then, if he agreed to the additional crop the sharecropping contract would be converted to a fixed rent. If he did not agree then the existing sharecropping arrangement would remain and no paddy would be planted. While not exactly the same environment, this arrangement resembles qualitative features of the contracts found in Models P1 and P2. In the good state, the contract provides less insurance to the agents so they work hard and in the other states it provides them with more insurance and does not make them work as hard. Evaluating the well suggests that the state is public information while talking to other tenants suggests that at least some information about the state is private.

The surveys indicate that communication is an important variable as theory predicts. Given the measurement issues involving implicit contracts, it is difficult to ascertain communication's precise role. Interestingly, the contracts with the non-resident landowners all were fixed rent contracts. These contracts make risk sharing over interim shocks more difficult, as suggested by the ex post disputes over rental payments. This also suggests that communication after the output has been realized might be important. In particular, communication between the landowner and other members of the village, including the group's financier, might be an important source of information. At this point, there are not models of this type of communication or of communication between agents and multiple principals, as in the case of a group with a landowner and a financier.

6 Conclusion

This paper investigated the effects of communication on incentives in a multi-stage optimal contracting model. The revelation principle was proven for the no-communication model. Linear programs for solving the models were derived and used to compute solutions. An alternative formulation for the communication model was provided. An advantage of the linear programming methods is that it can be used to study problems with arbitrary specification of preferences and endowments.

In simple examples, conditions for communication to be valuable were shown to depend on the type of equilibrium in the no-communication model. If this bindingness condition on incentive constraints was satisfied then Pareto-superior communication contracts which improved consumption risk sharing could be constructed. For a more complicated example communication not only changed consumption risk sharing but altered labor allocations. Action and consumption lotteries were also found to be valuable.

Finally, data from field work was reported. Interim shocks to production and labor endowments were reported. The precise role of communication was difficult to ascertain because of apparent implicit contingencies in the contracts. Nevertheless, communication appears to be an important variable, as predicted by theory.

A Proofs

The no-communication model forbids the agent from sending reports. The principal may still send messages $m \in M$, but these cannot be conditioned on anything. As in Model P2, the shock θ is hidden information. For a general proof of the revelation principle for games with no communication, see Forges (1986).

Definition 5 *A Bayesian-Nash equilibrium for Model P3 is a randomizer $\Delta(m)$, a consumption schedule $\Delta(c|q, m)$, and strategy $\mu^*(a|m, \theta)$ such that given the randomizer, and the consumption schedule, the strategy solves for each m, θ realization*

$$\max_{\mu(a|m, \theta)} \prod_{c, q, a} \Delta(c|q, m) p(q|a) \mu(a|m, \theta) U(c, \theta, a) \quad (19)$$

It is convenient to describe an object: the probability of consumption, output, and the strategy taken. Let a_Θ be the vector of state-dependent actions, which is referred to as a strategy. The probability of any strategy \mathbf{a}_Θ being taken by an agent is then:

$$\pi(\mathbf{a}_\Theta) = \prod_m \left(\prod_\theta \mu^*(a_\theta|m, \theta) \right) \Delta(m), \quad (20)$$

and the probability of consumption c , given output q , and strategy \mathbf{a}_Θ is

$$\begin{aligned} \pi(c|q, \mathbf{a}_\Theta) &= \prod_m \Delta(c|q, m) pr(m|\mathbf{a}_\Theta) \\ &= \prod_m \Delta(c|q, m) \frac{\left(\prod_\theta \mu^*(a_\theta|m, \theta) \right) \Delta(m)}{\pi(\mathbf{a}_\Theta)}, \end{aligned} \quad (21)$$

with

$$\pi(c, q, \mathbf{a}_\Theta) = \pi(c|q, \mathbf{a}_\Theta) \prod_\theta pr(\theta) p(q|a_\theta) \pi(\mathbf{a}_\Theta). \quad (22)$$

This determines an equilibrium distribution of consumption, output, and state contingent strategy. Before proving the revelation principle it is helpful to repeat the definition of the obedient game. Recall that the message space for this game is $M = \prod_{\theta \in \Theta} A$.

Definition 3 *An incentive-compatible allocation for Model P3 is a randomizer $\pi(\mathbf{a}_\Theta)$, a consumption schedule $\pi(c|q, \mathbf{a}_\Theta)$, and a deterministic strategy $a_\theta = a^*(\mathbf{a}_\Theta, \theta)$ such that given the randomizer, the consumption schedule, and the recommended strategy \mathbf{a}_Θ , the strategy $a_\theta = a^*(\mathbf{a}_\Theta, \theta)$ satisfies*

$$\prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) p(q|a_\theta) U(c, \theta, a_\theta) \geq \prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) p(q|\hat{a}) U(c, \theta, \hat{a}), \quad \forall \hat{a} \in A, \quad (23)$$

for each realization of θ .

Now it will be shown that the two mechanisms deliver the same allocations.

Theorem 1 *The original mechanism and the obedient game for Model P3 deliver the same allocations.*

Proof: Equivalence is shown by demonstrating that any allocation feasible under one kind is feasible under the other. Take any equilibrium satisfying the original mechanism. Construct the obedient game using equations (20) and (21). Assume that this equilibrium is not incentive compatible in the truth-telling game for some strategy \mathbf{a}_Θ , and realization $\tilde{\theta}$. Pick any better action and call it \hat{a} then:

$$\prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) p(q|a_{\tilde{\theta}}) U(c, \tilde{\theta}, a_{\tilde{\theta}}) < \prod_{c,q} \pi(c|q, \mathbf{a}_\Theta) p(q|\hat{a}) U(c, \tilde{\theta}, \hat{a}). \quad (24)$$

Substitute in equation (21) to get

$$\prod_{c,q} \prod_m \Delta(c|q, m) \frac{\left(\prod_{\theta} \mu^*(a_{\tilde{\theta}}|m, \theta) \Delta(m) \right)}{\pi(\mathbf{a}_\Theta)} p(q|a_{\tilde{\theta}}) U(c, \tilde{\theta}, a_{\tilde{\theta}}) \quad (25)$$

$$< \prod_{c,q} \prod_m \Delta(c|q, m) \frac{\left(\prod_{\theta} \mu^*(a_{\tilde{\theta}}|m, \theta) \Delta(m) \right)}{\pi(\mathbf{a}_\Theta)} p(q|\hat{a}) U(c, \tilde{\theta}, \hat{a}). \quad (26)$$

Now eliminate $\pi(\mathbf{a}_\Theta)$ from both sides. Next, observe since the above equation does not hold for the sum of $m \in M$, it must not hold for at least one m . Taking one of these m and calling it \tilde{m} means

$$\prod_{c,q} \Delta(c|q, \tilde{m}) \left(\prod_{\theta} \mu^*(a|\tilde{m}, \theta) \Delta(\tilde{m}) \right) p(q|a_{\tilde{\theta}}) U(c, \tilde{\theta}, a_{\tilde{\theta}}) < \prod_{c,q} \Delta(c|q, \tilde{m}) \left(\prod_{\theta} \mu^*(a|\tilde{m}, \theta) \Delta(\tilde{m}) \right) p(q|\hat{a}) U(c, \tilde{\theta}, \hat{a}). \quad (27)$$

Next construct a new strategy $\tilde{\mu}(\hat{a}|\tilde{m}, \tilde{\theta}) = \mu^*(\hat{a}|\tilde{m}, \tilde{\theta}) + \mu^*(\tilde{a}|\tilde{m}, \tilde{\theta})$, $\tilde{\mu}(\tilde{a}|\tilde{m}, \tilde{\theta}) = 0$ and let $\tilde{\mu}(a|m, \theta) = \mu^*(a|m, \theta)$ if any one of the following conditions hold: $\theta \neq \tilde{\theta}, m \neq \tilde{m}, a \neq \hat{a}, \tilde{a}$. Divide both sides of the previous equation by the common factor $(\prod_{\theta \neq \tilde{\theta}} \mu^*(a|\tilde{m}, \theta) \Delta(\tilde{m}))$, and then sum over all a to get the equation

$$\prod_{c,q,a} \Delta(c|q, m) \mu^*(a|\tilde{m}, \theta) p(q|a_{\tilde{\theta}}) U(c, \tilde{\theta}, a_{\tilde{\theta}}) \quad (28)$$

$$< \prod_{c,q,a} \Delta(c|q, m) \tilde{\mu}(a|\tilde{m}, \theta) p(q|a) U(c, \tilde{\theta}, a) \quad (29)$$

This contradicts our assumption that $\mu^*(a|m, \theta)$ was incentive compatible. So the obedient game which leaves the allocations unchanged is feasible. Going in the other direction is trivial, just set $M = \prod_{\theta} A$. Q.E.D.

The next proposition is proven for preferences of the form $U(c, \theta, a) = U(c) + V(\theta, a)$, with U strictly concave, $V_a < 0$, $V_{aa} < 0$, and $V_{\theta a} > 0$. The proof is easily modified for the preferences used in the many-action example.

Proposition 2 *Assume that the preferences above hold, that there are two outputs with the probability of the high output increasing with the action, and that in the solution to the no-communication model the incentive constraint does not bind on the low- θ agent. Then, communication is valuable.*

Proof: For simplicity, assume that consumption is a continuum. With concave utility and separable preferences there will not be consumption lotteries in the solution. Let a_l and a_h be the actions taken by θ_l and θ_h type, respectively, in the solution to the no-communication model solution. The optimal contract $c(q)$ satisfies

$$\prod p(q|a_h) U(c(q)) + V(\theta_h, a_h) \geq \prod p(q|a_l) U(c(q)) + V(\theta_h, a_l), \quad (30)$$

$$\prod_q p(q|a_l) U(c(q)) + V(\theta_l, a_l) > \prod_q p(q|\hat{a}) U(c(q)) + V(\theta_l, \hat{a}), \quad \forall \hat{a}. \quad (31)$$

Now add communication. Communication contracts are indexed by the state, that is, $c(q, \theta_h)$ and $c(q, \theta_l)$. Let $c(q, \theta_h) = c(q)$ and choose $c(q, \theta_l)$ such that

$$\times_q p(q|a_l)U(c(q, \theta_l)) = \times_q p(q|a_l)U(c(q)), \quad (32)$$

and

$$\times_q p(q|a_l)U(c(q, \theta_l)) + V(\theta_l, a_l) = \times_q p(q|a)U(c(q, \theta_l)) + V(\theta_l, a), \quad (33)$$

where $a < a_l$ is the downward adjacent action to a_l . This contract is constructed by increasing $c(q_l, \theta)$ and lowering $c(q_h, \theta)$ until (33) holds. This new contract is Pareto superior because it leaves the agent's utility unchanged and because $\mathbb{P}_q p(q|a_h)c(q, \theta_l) < \mathbb{P}_q p(q|a_h)c(q)$ is better for the principal. It still remains to show that the new contract preserves incentive compatibility.

The θ_l agent has no incentive to report θ_h since $c(q, \theta_h) = c(q)$. By the way $c(q, \theta_l)$ was constructed, $\mathbb{P}_q p(q|\tilde{a})U(c(q, \theta_l)) < \mathbb{P}_q p(q|\tilde{a})U(c(q))$, so the θ_l agent will not choose an action $\tilde{a} > a_l$. This same agent will also not choose an action $\tilde{a} < a_l$ because of the way the new contract was constructed.

For the θ_h agent, if he reports truthfully then a_h is incentive compatible because his contract has not changed from that of the no-communication model. It only remains to show that he will tell the truth.

Note that because of (30) and (32),

$$\times_q p(q|a_h)U(c(q, \theta_h)) + V(\theta_h, a_h) \geq \times_q p(q|a_l)U(c(q, \theta_l)) + V(\theta_h, a_l),$$

that is, the θ_h agent will not lie and choose action a_l . Now consider the strategy of lying and choosing $\tilde{a} < a_l$. Because the θ_l agent will not choose an action $\tilde{a} < a_l$ and since $V_{\theta_a} > 0$

$$\times_q p(q|a_l)U(c(q, \theta_l)) + V(\theta_h, a_l) \geq \times_q p(q|\tilde{a})U(c(q, \theta_l)) + V(\theta_h, \tilde{a}), \quad \forall \tilde{a} < a_l.$$

But by (30) and (32), this means that

$$\times_q p(q|a_h)U(c(q, \theta_h)) + V(\theta_h, a_h) \geq \times_q p(q|\tilde{a})U(c(q, \theta_l)) + V(\theta_h, \tilde{a}), \quad \forall \tilde{a} < a_l.$$

Finally, consider the strategy of the θ_h type lying and taking $\tilde{a} > a$. By equation (32), the construction that $c(q_l, \theta_h) < c(q_l, \theta_l) < c(q_h, \theta_l) < c(q_h, \theta_h)$, and the probability of the high output increasing with the action implies that $\times_q p(q|\tilde{a})(U(c(q, \theta_h)) - U(c(q, \theta_l))) > 0$. Therefore,

$$\begin{aligned} \times_q p(q|\tilde{a})U(c(q, \theta_l)) + V(\theta_h, \tilde{a}) &< \times_q p(q|\tilde{a})U(c(q, \theta_h)) + V(\theta_h, \tilde{a}) \\ &\leq \times_q p(q|a_h)U(c(q, \theta_h)) + V(\theta_h, a_h), \end{aligned}$$

with the latter inequality by the assumed incentive compatibility of the no-communication contract and the construction that $c(q) = c(q, \theta_h)$. Q.E.D.

Proposition 3 proves that if preferences are of the form $U(c, \theta, a) = U(c) + V(\theta, a)$, with U strictly concave, $V_a < 0$, $V_{aa} < 0$, and $V_{\theta a} > 0$ then many of the strategies in Model P3, the no-communication model, are not incentive compatible. When solving examples, this allows a sizable reduction in the grid space for that model. The Proposition's conditions directly apply to the example with the $V(\theta, a) = V(\theta - a)$ preferences. To apply the Proposition to the $V(\theta, a) = -\theta a^2$ preferences, the condition needs to be changed to $V_{\theta a} < 0$ because the low value of θ is now the low marginal disutility of effort agent.

Proposition 3 *For the preferences described above, in Model P3 strategies where the low- θ agent takes an action higher than the high- θ agent are not incentive compatible.*

Proof: As before, for simplicity assume that C is a continuum so that attention can be restricted to deterministic compensation schedules. Let $(a(\theta_l), a(\theta_h))$ be a pair of actions to implement. The incentive constraints include

$$\times_q p(q|a(\theta_l))U(c(q)) + V(\theta_l, a(\theta_l)) \geq \times_q p(q|a(\theta_h))U(c(q)) + V(\theta_l, a(\theta_h)),$$

$$\times \int_q p(q|a(\theta_h))U(c(q)) + V(\theta_h, a(\theta_h)) \geq \times \int_q p(q|a(\theta_l))U(c(q)) + V(\theta_h, a(\theta_l)).$$

Rearranging the two incentive constraints obtains

$$V(\theta_l, a(\theta_h)) - V(\theta_l, a(\theta_l)) \leq V(\theta_h, a(\theta_h)) - V(\theta_h, a(\theta_l)). \quad (34)$$

Now assume that $a(\theta_l) > a(\theta_h)$. This assumption and inequality (34) contradict the assumption that $V_{\theta a} > 0$. Q.E.D.

References

- [1] Baiman, Stanley, and John H. Evans III. "Pre-Decision Information and Participative Management Control Systems." *Journal of Accounting Research* 21 (Autumn 1983): 371-95.
- [2] Christensen, John. "Communication in Agencies." *Bell Journal of Economics* (Autumn 1981): 661-74.
- [3] Cole, Harold L. "Comment: General Competitive Analysis in an Economy with Private Information." *International Economic Review*, vol. 30 (February 1989): 249-52.
- [4] Demougin, Dominique M. "A Renegotiation-proof Mechanism for a Principal-agent Model with Moral Hazard and Adverse Selection." *RAND Journal of Economics* 20 (Summer 1989): 256-67.
- [5] Forges, Françoise. "An Approach to Communication Equilibria." *Econometrica* 54 (November 1986): 1375-85.
- [6] Hart, Oliver and Bengt Holmström. "The Theory of Contracts." In *Advances in Economic Theory: Fifth World Congress*. (Truman F. Bewley Ed.) Cambridge, England: Cambridge University Press, 1987: 71-155.
- [7] Holmström, Bengt. "Moral Hazard and Observability." *Bell Journal of Economics* (Spring 1979): 74-91.
- [8] Laffont, Jean-Jacques, and Jean Tirole. "Using Cost Observation to Regulate Firms." *Journal of Political Economy* 94 (June 1986): 614-41.
- [9] McAfee, R. Preston and John McMillan. "Competition for Agency Contracts." *RAND Journal of Economics* 18 (Summer 1982): 296-307.

- [10] Melumad, Nahum D. and Stefan Reichelstein. “Value of Communication in Agencies.” *Journal of Economic Theory* 47 (April 1989): 334-68.
- [11] Myerson, Roger B. “Optimal Coordination Mechanisms in Generalized Principal-Agent Problems.” *Journal of Mathematical Economics* 10 (June 1982): 67-81.
- [12] Myerson, Roger B. “Multistage Games With Communication.” *Econometrica* 54 (March 1986): 323-58.
- [13] Penno, Mark. “Asymmetry of Pre-Decision Information and Managerial Accounting.” *Journal of Accounting Research* 22 (Spring 1984): 177-91.
- [14] Prescott, Edward C. and Robert M. Townsend. “Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard.” *Econometrica* 52 (January 1984): 21-54.
- [15] Prescott, Edward S. “Communication in Models with Private Information: Theory, Applications, and Computation.” Ph.D. Dissertation. University of Chicago, August, 1995.
- [16] Prescott, Edward S. “Computing Solutions to Moral Hazard Programs Using the Dantzig-Wolfe Decomposition Algorithm.” Manuscript 2000.
- [17] Prescott, Edward S. and Robert M. Townsend. “The Boundaries and Collectiveness of Economic Organizations.” Manuscript 1999.
- [18] Rochet, Jean-Charles. “Solvency Regulations and the Management of Banking Risks.” *European Economic Review* 43 (April 1999): 981-90.
- [19] Suh, Yoon S. “Communication and Income Smoothing Through Accounting Method Choice.” *Management Science* 36 (June 1990): 704-23.

- [20] Townsend, Robert M. and Rolf A. E. Mueller “Mechanism Design and Village Economies: From Credit to Tenancy to Cropping Groups.” *Review of Economic Dynamics* 1 (January 1998): 119-72.
- [21] Townsend, Robert M. *The Medieval Village Economy: A Study of the Pareto Mapping in General Equilibrium Models*. Princeton, N.J.: Princeton University Press, 1993.
- [22] Walker, T.S. and James G. Ryan. *Village and Household Economies in India's Semi-Arid Tropics*. Baltimore: Johns Hopkins University Press, 1990.